DOCUMENT RESUME

ED 070 020 CG 007 706

AUTHOR	Harris, Margaret L.; Harris, Chester W.
TITLE	Item Analyses and Reliabilities for Reference Tests for Cognitive Abilities: Fifth Grade Boys and Girls.
INSTITUTION	Wisconsin Univ., Madison. Research and Development Center for Cognitive Learning.
REPORT NO	WRDCCL-TR-191
BUREAU NO	BR-5-0216
PUB DATE	Nov 71
CONTRACT	OEC-5-10-154
NOTE	26p.
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Cognitive Ability; Grade 5; *Item Analysis; *Test Construction; Testing; Tests; Test Validity

ABSTRACT

Test development efforts for constructing 35 of the tests included in a battery of 56 tests of possible cognitive abilities are described. Data were collected on 172 boys and 210 girls who had just completed the fifth grade. Item and total score statistics obtained for the 56 tests in the total battery are presented and discussed. (Author)

FILMED FROM BEST AVAILABLE COPY

ED 070020

TECHNICAL REPORT NO. 191

ITEM ANALYSES AND RELIABILITIES FOR REFERENCE TESTS FOR COGNITIVE ABILITIES: FIFTH GRADE BOYS AND GIRLS

REPORT FROM THE PROJECT ON A STRUCTURE OF CONCEPT ATTAINMENT ABILITIES

CG 007 706

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

WISCONSIN RESEARCH AND DEVELOPMENT
CENTER FOR
COGNITIVE LEARNING

THE UNIVERSITY OF WISCONSIN
Madison, Wisconsin

U. S. Office of Education
Center No. C-03
Contract OE 5-10-154

1

RESEARCH
COGNITIVE
LEARNING
DEVELOPMENT

ED 070020

Technical Report No. 191

ITEM ANALYSES AND RELIABILITIES FOR REFERENCE TESTS FOR COGNITIVE ABILITIES: FIFTH GRADE BOYS AND GIRLS

By Margaret L. Harris and Chester W. Harris

Report from the Project on
A Structure of Concept Attainment Abilities

Robert E. Davidson, Lester S. Golub, Herbert J. Klausmeier,
Thomas A. Romberg, B. Robert Tabachnick, Alan M. Voelker
Principal Investigators

Mary R. Quilling, Director
The Quality Verification Program

Wisconsin Research and Development
Center for Cognitive Learning
The University of Wisconsin
Madison, Wisconsin

November 1971

Published by the Wisconsin Research and Development Center for Cognitive Learning, supported in part as a research and development center by funds from the United States Office of Education, Department of Health, Education, and Welfare. The opinions expressed herein do not necessarily reflect the position or policy of the Office of Education and no official endorsement by the Office of Education should be inferred.

Center No. C-03 / Contract OE 5-10-154

Statement of Focus

The Wisconsin Research and Development Center for Cognitive Learning focuses on contributing to a better understanding of cognitive learning by children and youth and to the improvement of related educational practices. The strategy for research and development is comprehensive. It includes basic research to generate new knowledge about the conditions and processes of learning and about the processes of instruction, and the subsequent development of research-based instructional materials, many of which are designed for use by teachers and others for use by students. These materials are tested and refined in school settings. Throughout these operations behavioral scientists, curriculum experts, academic scholars, and school people interact, insuring that the results of Center activities are based soundly on knowledge of subject matter and cognitive learning and that they are applied to the improvement of educational practice.

This Technical Report is from the Quality Verification Program and from the Project on the Structure of Concept Attainment Abilities in Program 1. The Concept Attainment staff took primary initiative in identifying basic concepts in social studies at intermediate grade levels, while the Quality Verification Program assisted in developing tests to measure concept achievement and identifying reference tests for cognitive abilities. The tests will be used to study the relationships among cognitive abilities and learned concepts in various subject matter areas. The outcome of the Project will be a formulation of a model of structure of abilities in concept attainment in a number of subjects, including mathematics, science, and language arts, as well as social studies.

Contents

		Page
	List of Tables	vi
	Abstract	vii
I	Introduction	1
II	Procedures	5
	Test Development	5
	Test Revision	6
	Subjects	6
	Data Collection	8
	Treatment of the Data	8
III	Results and Discussion	11
	Reliability Estimates and Test Statistics	11
	Item Indices	11
IV	Summary and Conclusions	17
	Conclusions	17
	References	19
	Appendix A: Alphabetical Listing of Tests	21
	Appendix B: Source of Test and Number of Choices for Each Item	22

List of Tables

Table		Page
1	Classifications in Each of the Schemata	2
2	Test Data for Population and Samples	7
3	Distribution of Fathers' Occupations	9
4	Means, Standard Deviations, and Reliability Estimates: Boys and Girls	13
5	Item Indices	14

Abstract

Test development efforts for constructing 35 of the tests included in a battery of 56 tests of possible cognitive abilities are described. Data were collected on 172 boys and 210 girls who had just completed the fifth grade. Item and total score statistics obtained for the 56 tests in the total battery are presented and discussed.

I
Introduction

The primary objective of the project entitled "A Structure of Concept Attainment Abilities" (hereafter referred to as the CAA Project) is to formulate one or more models or structures of concept attainment abilities, and to assess their consistency with actual data. The major steps for attaining this primary objective were taken to be:

1. To identify basic concepts in language arts, mathematics, science, and social studies appropriate at the fourth grade level,
2. To develop tests to measure achievement of these concepts,
3. To identify reference tests for cognitive abilities, and
4. To study the relationships among learned concepts in these four subject matter fields and the identified cognitive abilities.

There are two major phases of Step 3. One is the examination of available systems for defining cognitive abilities followed by the selection and/or construction of tests implied by these systems. These efforts are described in "Three Systems of Classifying Cognitive Abilities as Bases for Reference Tests" (Harris & Harris, in press(b)). The present paper contains a description of the procedures used for constructing some of the tests implied by these systems and gives summary item and test statistics for all of the tests in the battery compiled as administered to two different samples, one composed of fifth grade boys and one of fifth grade girls. The second major phase of Step 3 is the empirical study of the interrelations of these tests in an attempt to validate and/or reconstruct these systems, and will be reported elsewhere.

Three fairly well-known systems for defining general cognitive abilities were analyzed to determine the nature of possible reference tests for cognitive abilities. They are the Guilford (1967) analysis of cognition using three contents and six products; the facet design for achievement, consisting of three tasks and three types of content, proposed by Guttman (1970); and the Primary Mental Abilities scheme of the Thurstones (1938, 1941). In analyzing these three systems it became apparent that modifications in the schemata of both Guilford and Guttman might be appropriate for content and the operation or task required. This led to a fourth schema for classifying abilities that deal with cognizing concepts. It involves classifying the nature of exemplars as things or relations; the content as verbal-semantic, picture-semantic, number-semantic, figural, number-symbolic, letter-symbolic, or word-form; and the task as classifying, excluding, or naming. Each of the 56 tests was classified in the content category; only those tests dealing with the cognition of concepts were classified in the other two categories. This schema is discussed in more detail in Harris and Harris (in press (b)).

A battery of 56 tests was developed to study the relationships among the Guilford, Guttman, and Thurstone schemata. For factor analysis, it is desirable to have at least two, and preferably three, tests to measure each hypothesized ability (each of the identified possible cognitive abilities). Tests were selected, adapted, or constructed as specific measures of the ability implied by a cell of interest in at least one of the schemata, including the newly proposed system for the cognition of concepts. Since Guilford's Structure of Intellect model is the most specific of the three schemata analyzed, most of the tests were initially chosen from his point of view. It should be pointed out here, however, that many of these tests are of the same type as those initially studied

Table 1. Classifications in Each of the Schemata

Test Name	Guilford	Guttman	Thurstone	Cognition of Concepts		
1 Picture Meaning	(CMU)[a]	A:P	V		P-M	
2 Verbal Classification	CMC	RI:V	I	T	V-M	C
3 Number Series	CSS	RI:N	I	R	N-M	C
4 Remembering Classes: Members	(MMC)	RA:V	M		V-M	
5 Number Class Extension	(CSC)	RI:N	I	T	N-M	C
6 Word Groups	CSC	RI	I	T	W-F	C
7 Remembering Classes: Members II	(MMC)	RA:V	M		V-M	
8 Disemvowelled Words	CSU	A:V	W		V-M	
9 Letter Grouping	CSC	RI	I	T	L-S	E
10 Circle Reasoning	CSS	RI:P	I	R	F	C
11 Figure Exclusion	CFC	RI:P	I	T	F	E
12 Seeing Trends	CSR	RI	I	R	W-F	C
13 Picture Classification	(CMC)	RI:P	I	T	P-M	C
14 Paragraph Comprehension	CMU	A:V	V		V-M	
15 Remember Classes: Names	MMC	RA:V	M		V-M	
16 Word Group Naming	(NMC)	RI:V	I	T	V-M	N
17 Gestalt Completion	CFU	RA:P	C_1		P-M	
18 Card Rotations	CFT	RA:P	S		F	
19 Spatial Relations	CFT	RA:P	S		F	
20 Verbal Exclusion	CMC	RI:V	I	T	V-M	E
21 Best Word Class	EMC	RI:V	I	T	V-M	N
22 Omelet	CSU	A:V	W		V-M	
23 Picture Group Naming	(NMC)	RI:P	I	T	P-M	N
24 Concealed Words	CFU	RA:P	C_1		V-M	
25 Perceptual Speed	EFU	RA:P	P		F	
26 Letter Triangle	CSS	RI	I	R	L-S	C
27 Letter Classification	(CSC)	RI	I	T	L-S	C
28 Picture Class Memory	MMC	RA:P	M		P-M	
29 Puzzles	EMI	A:V	D		V-M	
30 Spelling	CSU	A:V	W		V-M	
31 Picture Exclusion	(CMC)	RI:P	I	T	P-M	E
32 Sensitivity to Order	CMR	RI:V	I	R	V-M	C
33 Figure Analogies	CFR	RI:P	I	R	F	C
34 Scrambled Sentences	CMU	A:V	V		V-M	
35 Same-Opposite	(CMU)	RA:V	V		V-M	
36 Figure Matrix	CFR	RI:P	I	R	F	C
37 Remote Class Completion	(NMC)	RI:V	I	T	V-M	C
38 Number Exclusion	(CSC)	RI:N	I	T	N-S	E
39 Sentence Order	NMS	RA:V	V		V-M	
40 Vocabulary	CMU	A:V	V		V-M	
41 Word Relations	CSR	RI	I	R	W-F	C
42 Verbal Analogies	CMR	RI:V	V	R	V-M	C
43 Best Trend Name	EMR	RI:V	I	R	V-M	N
44 Picture Arrangement	NMS	RA:P	V		P-M	
45 Arithmetic Problems	MSI	A:N	N		N-M	
46 Identical Pictures	EFU	RA:P	P		F	
47 Picture Group Name Selection	(EMC)	RI:P	I	T	P-M	N
48 Number Classification	CSC	RI:N	I	T	N-S	C
49 Word Exclusion	(CSC)	RI	I	T	W-F	E
50 Number Relations	CSC	RI:N	I	R	N-M	E
51 Word Linkage	CMR	A:V	V		V-M	
52 Figure Classification	CFC	RI:P	I	T	F	C
53 Class Name Selection	EMC	RI:V	I	T	V-M	N
54 Necessary Arithmetic Operations	CMS	RA:N	D		V-M/N-M	
55 Verbal Analogies III	EMR	RI:V	V	R	V-M	C
56 Remembering Classes: Members III	(MMC)	RA:V	M		V-M	

[a] Parentheses indicate developer's identification and/or source.

Table 1 (Continued)

Key to Schema Classifications:

Guilford

Operations (first letter):
- C Cognition
- M Memory
- D Divergent Production
- N Convergent Production
- E Evaluation

Contents (second letter):
- M Semantic
- S Symbolic
- F Figural

Products (third letter):
- U Units
- C Classes
- R Relations
- S Systems
- T Transformations
- I Implications

Guttman

- RI Rule-Inferring
- RA Rule-Applying
- A Achievement or rule-applying when the rule used is formally taught in school
- V Verbal
- N Numerical
- P Pictorial

Thurstone

- S Spatial
- P Perceptual Speed
- N Numerical
- V Verbal
- W Word Fluency
- M Memory
- I Induction
- D Deduction
- C_1 Closure One

Cognition of Concepts

Nature of Exemplars (first column):
- T Things
- R Relations

Content (second column):
- V-M Verbal-Semantic
- P-M Picture-Semantic
- N-M Number-Semantic
- F Figural
- N-S Number-Symbolic
- L-S Letter-Symbolic
- W-F Word-Form

Task (third column):
- C Classify
- E Exclude
- N Name

by the Thurstones. Each of the tests can be classified, *a priori*, into a relevant cell of each of the schemata.

The nature of each of the 56 tests and their classification in each of the schemata is discussed in Harris and Harris (in press (b)). A summary of the classifications is given in Table 1. The tests are listed in the order of administration. (An alphabetical listing of the tests with numbers corresponding to Table 1 is given in Appendix A.) In the case of the Guilford classifications, those given in parentheses indicate our placement of the test in a particular cell in the Structure of Intellect; the others are Guilford's identifications. All of the other classifications are ours; it should be pointed out, however, that the basis for most of them was obtained from the literature.

Of the total battery of 56 tests, 35 were constructed for this project. These 35 include Concealed Words, Gestalt Completion, and Verbal Analogies of which some or all of the items were adapted from another source. These tests are described in "Newly Constructed Reference Tests for Cognitive Abilities" (Harris & Harris, in press (a)). The source of each of the 56 tests is given in Appendix B. It should be pointed out here that Tests 4, 7, and 56 are the same test given on three different occasions.

Test development procedures and results of data collection using the total battery will be discussed in the following sections.

II
Procedures

This section contains a discussion of the test development procedures used including initial item construction and revision of the tests based on item analysis results for the 35 tests that were constructed for the CAA Project. Data collection procedures, subjects, and treatment of the data for the total battery of 56 tests are discussed.

Test Development

Of the total battery of 56 tests, 35 were constructed by the researchers of the CAA Project specifically for use in the project, 19 are published tests (used in this project for research purposes, with permission from the various publishers), and the remaining 2 tests are merely further administrations of one of the memory tests. The test development procedures discussed in general terms here apply to the 35 tests that were constructed for this project. As was mentioned earlier, some items were adapted from another source.

The first step in the test construction process was the determination of the format and type of content for the items of each particular test. For example, for Verbal Classification the content was specified as verbal-semantic and the task as one of adding an exemplar to a set of four exemplars of a class. All examplars, both those defining the class and those offered as possible additional exemplars, were names of things rather than relationships.

When applicable, principles were enumerated by which items meeting the task and content specifications could be generated. As an example: various pathways that the letters could follow within a triangle; various sections of the alphabet; whether letters were to be adjacent as opposed to skipping one or skipping two; and the direction, forwards or backwards within the alphabet, were specified for the Letter Triangle test. When the enumeration of principles by which items could be generated was not applicable, as with Concealed Words, the advice of experts and small-scale tryouts were used to assist in determining appropriateness of specific items.

For all of the tests dealing with classes using semantic content (both verbal and pictorial), a sample of class names was drawn at random from a previously compiled long list of class names deemed appropriate for the fourth grade level. A sampled class name sometimes had to be dropped and replaced by another sampled one if the exemplars of that class could not be pictured or if there were not enough appropriate exemplars of the class.

The items for the Verbal Analogies test were taken from analogy questions given by Gouber (1967). He included 14 different relations and a number of different analogies for each of these. In the revised form of the test, we used two items each for 12 different types of relations. In so doing, we felt that the antonym and synonym relations were essentially the same for our purposes and that a grammatical relation was inappropriate.

Care was taken that all items for a test be of the same type or task. For example, for the Figure Classification test, items for which all of the exemplars and the correct choice formed a series were not used. An example of this would be each succeeding exemplar, and thus the correct answer, increasing in size. Care was taken also that all of the items be of the same type of content. For example, the Number Series test was designed to contain number-semantic content in a series-type format, using numbers as cardinal numbers. Items that could be answered correctly simply by form or symbolic content of the numbers were not used. An example of this is:

9 99 999 9999 ____

One could say that this item involves cardinality of the numbers—add 90, add 900, add 9000—but it can be answered correctly without conceptualizing this. If it were used, such an item would be placed in a series-type test using number-symbolic content.

An attempt was made to keep reading level and other possible confounding aspects at a minimum so that the items of the test could elicit the student behavior for which they were intended. For example, the Remembering Classes: Names test was intended to measure ability to remember inferred names for a number of classes of things. The class name first had to be inferred from the exemplars given, and then remembered. The exemplars, and hence class names, chosen for the items of this test were ones with which fifth graders were expected to be familiar.

The previous illustrations serve as examples of the kinds of things taken into consideration in the test construction process.

The items that were constructed were subjected to extensive critique during the writing process, after all had been completed, or at both stages. The advice of experts in a particular subject-matter field such as mathematics was utilized when developing tests involving number-semantic content and/or a particular area of interest such as memory (immediate versus longer term recall). The advice and critique of reading specialists was sought regarding vocabulary used and the appropriateness of the reading level of the items. Some or all items of many of the tests were tried out on a few fifth grade and/or other level subjects to assist in determining appropriateness of content and/or format. All items were subjected to a final critique by the writers of this paper.

Test Revision

Tryouts of the tests for item analysis and test revision purposes were conducted during the spring of 1969 and the winter and spring of 1970. The tryout version of each test consisted of at least 30 items; there were approximately 30 for most tests. Usually a half hour was allotted to the students to respond to the items for each of the tests. Fifteen of the tests were piloted with fourth grade students in the Hamilton School District of Sussex, Wisconsin, and the Hartford, Wisconsin, school district during late March and April of 1969. During the fall of 1970 the decision was made to conduct the entire CAA Project with fifth grade students instead of fourth grade students as had been originally planned. Thus, subsequent tryouts of cognitive abilities tests were conducted using fifth grade students as subjects and some minor revisions were made in some of the 15 tests that had been piloted using fourth grade students. Further pilot studies of cognitive abilities tests were conducted during January, 1970, with fifth grade students in the Hamilton School District and the West Allis, Wisconsin, school system. Fifth grade students in the Brookfield, Wisconsin, school system responded to a number of the tests during April, 1970. Approximately 100 students responded to each test. Most of the tests were piloted only once, though several were revised extensively following the first tryout and then were piloted again in the revised form.

The tryout data were subjected to the Generalized Item Analysis Program (GITAP) (Baker, 1969) the output of which provides the proportion responding, item-criterion biserial correlation, X_{50} (point on the criterion scale corresponding to the median of the item characteristic curve), and β (the reciprocal of the standard deviation of the item characteristic curve which is a measure of the discriminating power of the item) for each possible choice for each item as well as summary descriptive statistics for the total test. It also gives the Hoyt reliability for the total test and the standard error of measurement. Based upon both difficulty and discrimination indices obtained from the item analyses, usually 20 items were selected to be retained in a revised version of the test to be used in attaining Steps 3 and 4 of the project's primary objective as given on page 1. Items with the highest discrimination indices for the correct choice were selected with the stipulation that only items with X_{50}s within the range of -2.00 to +2.00 and, insofar as possible, well-representing that range be retained. If the β for one or more of the incorrect choices was an undesirable value, though otherwise the item was functioning well, those incorrect choices were revised.

The revised version of each of the 35 tests constructed for the CAA Project can be found in "Newly Constructed Reference Tests for Cognitive Abilities" (Harris & Harris, in press (a)).

Subjects

The battery of 56 cognitive abilities tests was administered during summer, 1970, to 172 boys and 210 girls who had just completed

the fifth grade in the public school system of Madison, Wisconsin. The students were randomly selected from the populations of all such girls and boys. The Madison Public School System made available the information concerning the populations and used their computing facilities to designate the random sample for the girls. Project researchers identified the random sample for the boys.

Initially, a random sample of 350 girls was drawn. Letters were sent to the parents of these students explaining the purpose and details of the testing, and inviting their daughter to participate in the testing program. A stamped and addressed postcard was enclosed which the parents were asked to complete and return indicating whether or not they were willing to allow their daughter to participate. One hundred and thirty-nine *yes* responses and 62 *no* responses were obtained from the cards returned. Those parents who had not returned the card by a specified date were phoned. An additional 49 *yes* and 72 *no* responses were obtained by phone. Since this total of *yes* responses did not give as many subjects as were desired, an additional sample of 100 girls was drawn at random. From this sample, 33 *yes* and 30 *no* responses were obtained by card. Thus, of the total sample of 450 girls, 221 *yes* and 164 *no* responses were received; 11 students did not complete the testing, resulting in a total of 210 girls tested. These students were paid $15.00 for participating.

A random sample of 450 boys was drawn and letters were sent. By mail, 136 *yes* and 34 *no* responses were obtained. An additional 36 *yes* and 80 *no* responses were obtained by phone. From an additional sample of 80 boys drawn at random, 21 *yes* and 45 *no* responses were obtained. Thus, of the total sample of 530 boys, 193 *yes* and 159 *no* responses were received; 21 students did not complete the testing which resulted in a total of 172 boys tested. As with the girls, the boys who completed the testing program were paid $15.00.

Since the participation of all students comprising the random samples was impossible to attain, test score and IQ data were obtained from the files of the Madison Public School System, for both the school population and

Table 2. Test Data for Population and Samples

Test		Population	Boys	Girls
Lorge-Thorndike Intelligence	$\bar{X}$	106.60	108.30	111.12
	s		15.05	13.82
	N	2605	157	206
Iowa Basic Skills				
Vocabulary	$\bar{X}$	5.53	5.69	5.66
	s		1.42	1.41
	N	2520	171	203
Reading Comprehension	$\bar{X}$	5.44	5.51	5.87
	s		1.55	1.42
	N	2520	171	203
Language Skills	$\bar{X}$	5.24	5.18	5.67
	s		1.44	1.32
	N	2520	171	202
Work-Study Skills	$\bar{X}$	5.46	5.71	5.73
	s		1.34	1.13
	N	2520	171	202
Arithmetic Skills	$\bar{X}$	5.05	5.24	5.24
	s		1.09	1.05
	N	2520	171	202
Composite	$\bar{X}$	5.35	5.46	5.64
	s		1.25	1.12
	N	2520	171	201

those participating students for whom the information was available. Table 2 includes the summary statistics for the population of fifth grade students in the Madison Public School System during the school year of 1969-70, and for the boys and the girls who comprised the tested samples for the cognitive abilities tests. The IQs were obtained in the fall of 1968, when the subjects were fourth graders, using the Lorge-Thorndike Intelligence Test; and the scores on the Iowa Tests of Basic Skills, given in grade equivalent scores, were obtained in the fall of 1969 when the subjects were fifth graders.

Data on fathers' occupations were collected from the students using the Master Occupational Code of the United States Bureau of the Census. These data were tabulated and are presented in Table 3.

Data Collection

The data for the girls were collected in two centrally-located schools, one on the East side and one on the West side of the city, during nine $2\frac{1}{2}$-hour daily sessions over a two-week period. Subjects could choose the weeks and the school in which they wanted to report for testing. A two-week session was held at Hawthorne School from June 29 to July 10 and a two-week session was held at Hoyt School from July 20 to July 31. Each $2\frac{1}{2}$-hour session consisted of the students responding to two booklets composed, in most cases, of three tests each, and a $\frac{1}{2}$-hour activity break after completing the first booklet.

The data for the boys were collected at the University of Wisconsin, during five $3\frac{1}{2}$-hour daily sessions for one week. Two different weekly sessions, August 17 to August 21 and August 24 to August 28, were held and the subjects could choose the week which they preferred.

The 56 tests in the battery were arranged in 18 booklets composed of three tests each, with the exception of two booklets that contained four tests each. The tests were given in the same order to all of the subjects in each of the samples. The students responded to the tests by marking their chosen response for each item directly on a machine-scorable answer sheet for 45 of the tests. For the remaining 11 tests, the students responded directly in the test booklet. The subjects' responses to these 11 tests were later coded onto machine-scorable answer sheets. All of the answer sheets were read by machine and the responses punched onto data cards. The tests were given by experienced test administrators to groups of approximately 30 subjects each.

The tests were not administered in a speeded fashion except for two tests, Perceptual Speed and Identical Pictures, which were designed to be speeded tests. A suggested time limit was given for each test. If five or more students were still working at the end of the time limit, the time for the test was extended until fewer than five students were still working. If all subjects finished the test before the suggested time limit was up, the test administrator went on to the next test. The suggested time limit was ample for every subject to complete most of the tests.

Treatment of the Data

The treatment of the data consisted of two main procedures: reliability estimation and item analysis. The data were analyzed separately for each sex group. Hoyt analysis of variance reliability estimates were obtained for each of the tests as well as means and standard deviations.

Item analyses using the GITAP program (Baker, 1969) were obtained for each of the items as a part of its total test score. This program provides proportion responding, item-criterion biserial correlation, X_{50}, and β statistics for each choice of each item. The proportion of students who respond correctly to an item is an index of the difficulty level of that item. The greater the value of the difficulty index, the easier the item. The biserial correlation coefficient is an index of the discriminating ability of the item choice. For these analyses the criterion ability used was the appropriate total test score. X_{50} is the point on the criterion scale, given in standard deviation units, corresponding to the median of the item characteristic curve. It is the score point at which subjects have a 50-50 chance of choosing that response. β is the reciprocal of the standard deviation of the item characteristic curve at the X_{50} point. It is an index of the discrimination power of the item.

Table 3. Distribution of Fathers' Occupations

Occupation	Boys	Girls
PROFESSIONAL, TECHNICAL, AND KINDRED WORKERS		
00. Accountant	6	4
01. Architect	2	1
02. Dentist	--	1
03. Engineer	3	10
04. Lawyer, Judge	4	1
05. Clergyman	3	2
06. Doctor	9	3
07. Nurse	--	--
08. Teacher, Professor	15	9
09. Other Professional	12	21
FARMER		
11. Farmer	--	--
MANAGERS, OFFICIALS, PROPRIETORS, EXCEPT FARM		
21. Owner of Business	3	5
22. Manager, Official	19	20
CLERICAL AND KINDRED WORKERS		
31. Bookkeeper	1	1
32. Receptionist	--	--
39. Other Clerical and Kindred Workers	2	8
SALES WORKERS		
49. Salesman	20	19
CRAFTSMEN, FOREMEN, AND KINDRED WORKERS (SKILLED WORKERS)		
51. Craftsmen, Skilled Worker	20	10
52. Foreman	4	1
53. Armed Services—Officer	2	1
54. Armed Services—Enlisted Man	2	--
OPERATIVES AND KINDRED WORKERS (SEMI-SKILLED WORKERS)		
61. Truck Driver	5	5
62. Operative in Factory	9	9
69. Other Operative and Kindred Workers	13	32
PRIVATE HOUSEHOLD AND SERVICE WORKERS		
71. Fireman	--	4
72. Policeman	2	1
73. Other Protective Service Worker	--	--
74. Practical Nurse, Nurses Aide	1	2
75. Private Household Workers	--	--
79. Other Service Workers	17	23
81. Non-Farm Laborer	1	--
82. Farm Laborer	--	--
91. Not presently in labor force	2	6
99. Not ascertained	6	12

III
Results and Discussion

The means, standard deviations, and Hoyt reliability estimates obtained for the summer, 1970, test battery data are presented separately for boys and girls, for each test. Also included in this section are a presentation and discussion of a summary of the item indices obtained for the correct choice for each of the tests.

Reliability Estimates and Test Statistics

Table 4 contains the means, standard deviations, and Hoyt reliability estimates obtained for the data collected during summer, 1970, using the 56 tests in the battery. These 56 tests consist of 35 which were constructed specifically for the CAA Project, 19 published tests, and two which are merely two further administrations of one of the memory tests. The source of each test is given in Appendix B. The data were analyzed separately for the 172 boys and the 210 girls. Table 4 includes the results for both of the samples. The number of items composing each of the tests is given in Table 4. If one is interested, the number of choices for each item of each test can be found in Appendix B. The tests are included in Table 4 in the order of administration. An alphabetical listing of the tests is given in Appendix A.

The mean scores and standard deviations are very similar for boys and for girls. The reliability estimates are generally slightly higher for boys than they are for girls. The reliability estimates are, in general, quite good with only 13 of the 112 estimates below .70; 56 of the estimates are equal to or greater than .80 with 13 of these being equal to or greater than .90. Of the 13 reliability estimates below .70, seven of them are for tests using pictures as semantic content. Perhaps it is more difficult to build reliable tests using pictures for the stimulus material. The reliability estimates for one test, Verbal Analogies III, were quite low for both boys and girls. Evidently this test was too difficult for these subjects.

These are relatively short tests, but only numbers 25 (Perceptual Speed) and 46 (Identical Pictures) were administered in a speeded fashion. For these two speeded tests, the Hoyt reliability estimate probably is an overestimate. It is interesting to note that the reliability estimates for the memory test that was given on three different occasions—at the beginning of an hourly test session; at the end of this same hour session without restudy of the material; and as the last test in the battery, again without restudy of the material—remained almost identically the same over the three occasions. The means are much the same for the first two occasions but dropped somewhat for the third occasion. These three test administrations are numbered 4, 7, and 56.

The reliability estimates are sufficiently high to warrant study of the dimensionality of these selected cognitive abilities tests (with the exception of the Verbal Analogies III test), which is a major objective of the CAA Project and is the main purpose for developing this battery of 56 tests. The factor analyses of these data will be reported in a later paper.

Item Indices

Table 5 contains a summary by sex of the item indices obtained for each of the 56 tests. The indices included are proportion correct (this often is called difficulty or P), item-criterion biserial correlation, X_{50}, and β. They are given for the correct choice only. For proportion correct and biserial correlation, the highest, lowest, and mid-value obtained over all of the items for a test is given; for X_{50} and β only the mid-value over all of the

Table 4. Means, Standard Deviations, and Reliability Estimates: Boys and Girls

Test	Number of Items	Mean		Standard Deviation		Hoyt Reliability	
		Boys[a]	Girls[a]	Boys	Girls	Boys	Girls
1 Picture Meaning	30	21.85	20.74	4.28	4.66	.78	.80
2 Verbal Classification	20	16.13	16.53	3.34	3.02	.77	.74
3 Number Series	20	13.62	13.25	4.43	4.17	.85	.82
4 Remembering Classes: Members	20	14.49	15.08	3.60	3.58	.75	.77
5 Number Class Extension	20	10.40	9.72	4.74	3.98	.83	.74
6 Word Groups	20	14.85	15.31	4.09	3.73	.82	.80
7 Remembering Classes: Members II	20	14.26	15.11	3.90	3.51	.78	.75
8 Disemvowelled Words	32	21.66	23.45	7.82	6.42	.93	.89
9 Letter Grouping	20	11.52	12.50	3.95	3.48	.77	.72
10 Circle Reasoning	20	8.43	9.90	4.94	4.61	.86	.83
11 Figure Exclusion	25	20.06	20.00	2.89	2.70	.64	.59
12 Seeing Trends	20	11.91	12.52	4.30	4.32	.80	.81
13 Picture Classification	20	12.74	12.53	3.28	2.84	.69	.59
14 Paragraph Comprehension	32	18.23	19.45	6.27	5.60	.84	.80
15 Remembering Classes: Names	20	16.96	18.36	3.28	2.39	.81	.79
16 Word Group Naming	20	12.25	12.30	3.73	3.40	.75	.68
17 Gestalt Completion	20	12.81	12.85	3.43	3.67	.72	.76
18 Card Rotations	112	92.02	86.38	22.34	20.84	.98	.97
19 Spatial Relations	25	16.70	16.75	4.58	4.31	.81	.78
20 Verbal Exclusion	20	13.42	13.80	3.23	2.61	.70	.56
21 Best Word Class	20	12.35	12.61	4.11	3.78	.79	.75
22 Omelet	20	10.87	12.21	4.44	4.21	.84	.82
23 Picture Group Naming	20	12.74	12.02	3.54	3.37	.72	.67
24 Concealed Words	20	9.77	8.67	3.88	3.45	.78	.72
25 Perceptual Speed	40	21.76	23.06	6.32	5.44	.89	.86
26 Letter Triangle	20	14.78	15.45	3.68	3.64	.79	.81
27 Letter Classification	20	12.72	13.87	4.30	3.77	.81	.78
28 Picture Class Memory	20	15.33	15.60	2.98	3.29	.69	.76
29 Puzzles	30	22.09	23.41	4.39	4.24	.74	.76
30 Spelling	30	17.14	19.45	7.00	6.32	.89	.87
31 Picture Exclusion	20	12.74	12.45	2.92	2.92	.56	.57
32 Sensitivity to Order	20	13.63	13.94	4.26	3.98	.84	.81
33 Figure Analogies	22	15.89	16.64	5.95	4.87	.92	.88
34 Scrambled Sentences	20	13.47	13.45	2.88	2.87	.53	.52
35 Same-Opposite	20	16.27	16.50	3.86	3.51	.84	.81
36 Figure Matrix	20	10.95	9.36	4.63	4.11	.82	.76
37 Remote Class Completion	26	12.73	13.08	4.32	4.42	.77	.80
38 Number Exclusion	20	11.97	12.15	4.30	3.53	.81	.71
39 Sentence Order	40	32.42	34.31	8.01	7.01	.93	.93
40 Vocabulary	30	19.24	19.20	6.70	6.00	.89	.86
41 Word Relations	20	12.66	13.07	5.44	5.64	.90	.91
42 Verbal Analogies	24	14.92	14.82	4.49	4.69	.78	.81
43 Best Trend Name	20	10.65	10.50	4.12	3.71	.77	.70
44 Picture Arrangement	16	9.86	9.61	2.37	2.71	.50	.63
45 Arithmetic Problems	35	16.58	16.59	7.19	7.40	.90	.91
46 Identical Pictures	48	22.09	25.44	5.39	6.17	.88	.91
47 Picture Group Name Selection	20	14.41	14.00	3.01	3.21	.66	.70
48 Number Classification	30	22.75	22.44	7.24	7.31	.93	.93
49 Word Exclusion	20	11.67	11.33	3.70	3.45	.73	.68
50 Number Relations	20	11.94	10.37	5.16	4.52	.87	.81
51 Word Linkage	20	11.48	11.90	4.28	4.34	.78	.79
52 Figure Classification	20	14.42	14.43	3.84	4.13	.82	.82
53 Class Name Selection	20	15.36	14.94	3.59	3.74	.80	.81
54 Necessary Arithmetic Operations	15	10.53	10.48	3.78	3.74	.85	.84
55 Verbal Analogies III	20	6.44	6.15	2.43	2.44	.39	.43
56 Remembering Classes: Members III	20	13.05	13.72	3.88	3.71	.77	.75

[a]Number of subjects is 172 for boys and 210 for girls.

Table 5. Item Indices

Test[b]	Proportion Correct[a]						Biserial Correlation[a]						X_{50}		β	
	High		Low		Mid		High		Low		Mid		Mid		Mid	
	B	G	B	G	B	G	B	G	B	G	B	G	B	G	B	G
1 Picture Meaning	99	100	10	16	77	66	88	79	14	05	61	57	-1.09	- .71	.76	.68
2 Verbal Classification	95	97	70	61	82	84	126	128	42	37	64	69	-1.30	-1.46	.83	.94
3 Number Series	94	93	39	30	67	67	93	86	43	44	70	66	- .66	- .72	.99	.88
4 Remembering Classes: Members	90	89	42	46	75	76	89	83	24	01	58	68	-1.21	-1.15	.71	.92
5 Number Class Extension	72	69	28	28	50	47	73	63	46	31	62	54	.01	.13	.80	.64
6 Word Groups	92	95	44	40	74	80	90	99	35	33	72	71	-1.02	-1.18	1.04	1.00
7 Remembering Classes: Members II	82	90	39	50	75	75	87	82	33	06	58	61	- .94	-1.03	.72	.77
8 Disemvowelled Words	95	98	37	44	69	72	101	102	45	32	78	73	- .59	- .82	1.26	1.07
9 Letter Grouping	80	87	18	21	61	65	91	98	16	12	62	62	- .43	- .64	.79	.80
10 Circle Reasoning	69	73	18	22	42	52	92	86	25	32	72	64	.27	- .06	1.03	.84
11 Figure Exclusion	99	99	40	35	87	87	110	153	06	26	55	51	-1.66	-1.96	.67	.59
12 Seeing Trends	87	76	20	22	64	67	88	94	-08	-03	66	65	- .49	- .71	.88	.86
13 Picture Classification	91	96	28	11	67	69	74	70	29	03	54	50	- .73	- .91	.64	.57
14 Paragraph Comprehension	87	85	32	37	56	59	82	70	34	17	52	51	- .25	- .57	.61	.60
15 Remembering Classes: Names	97	98	68	74	87	93	100	126	39	53	70	82	-1.39	-1.80	.99	1.45
16 Word Group Naming	95	94	10	16	62	65	78	71	24	16	56	51	- .62	- .80	.67	.59
17 Gestalt Completion	99	99	28	28	65	65	104	75	47	36	57	62	- .81	- .70	.69	.78
18 Card Rotations	93	92	62	46	82	78	114	98	37	21	80	68	-1.19	-1.16	1.35	.93
19 Spatial Relations	90	94	20	23	71	70	83	68	35	39	57	58	- .83	- .80	.70	.87
20 Verbal Exclusion	98	97	05	02	71	74	117	84	18	18	57	48	-1.01	-1.12	.69	.55
21 Best Word Class	88	94	26	21	67	65	78	87	10	-08	61	59	- .63	- .65	.77	.74
22 Omelet	91	91	06	13	56	66	77	83	51	46	69	65	- .23	- .60	.94	.86
23 Picture Group Naming	98	97	36	33	58	55	74	65	07	10	54	52	- .36	- .27	.64	.61
24 Concealed Words	79	83	06	04	50	39	78	74	35	20	61	57	- .01	.41	.77	.69
25 Perceptual Speed	98	100	03	03	70	76	104	95	20	-24	77	68	- .67	-1.11	1.21	.93
26 Letter Triangle	91	94	42	45	79	82	82	90	47	49	63	68	-1.15	-1.37	.81	.93
27 Letter Classification	90	93	39	36	66	68	80	88	37	36	63	62	- .66	- .73	.82	.80
28 Picture Class Memory	98	96	35	38	81	80	95	102	27	-23	69	76	-1.03	- .96	.94	1.18
29 Puzzles	92	93	41	48	76	79	68	71	17	16	48	54	-1.34	-1.69	.54	.65
30 Spelling	81	89	24	36	60	66	80	80	35	30	63	63	- .37	- .67	.80	.81

Table 5 (Continued)

Test[b]	Proportion Correct[a]						Biserial Correlation[a]						X_{50}		β	
	High		Low		Mid		High		Low		Mid		Mid		Mid	
	B	G	B	G	B	G	B	G	B	G	B	G	B	G	B	G
31 Picture Exclusion	96	96	33	22	63	64	65	66	27	34	43	43	- .78	- .72	.47	.48
32 Sensitivity to Order	88	90	19	21	76	76	95	95	04	18	74	67	- .79	- .93	1.11	.90
33 Figure Analogies	84	91	57	58	74	78	116	115	47	47	82	77	- .72	-1.00	1.45	1.21
34 Scrambled Sentences	94	87	45	47	68	64	76	59	23	16	40	43	- .94	- .85	.43	.47
35 Same-Opposite	88	94	65	69	84	81	96	98	45	45	79	73	-1.22	-1.33	1.28	1.08
36 Figure Matrix	74	81	38	26	51	46	83	71	21	26	65	55	- .07	.22	.86	.65
37 Remote Class Completion	92	94	12	05	51	45	77	70	20	22	55	61	- .05	.20	.65	.77
38 Number Exclusion	94	99	37	35	60	59	95	76	33	-01	63	58	- .43	- .45	.81	.70
39 Sentence Order	94	95	61	60	83	89	102	125	38	42	80	82	-1.18	-1.36	1.35	1.44
40 Vocabulary	87	90	33	30	60	63	91	87	32	23	66	58	- .46	- .63	.88	.71
41 Word Relations	78	84	26	39	64	66	88	96	51	53	80	84	- .44	- .53	1.34	1.58
42 Verbal Analogies	89	89	24	26	68	64	86	90	24	35	58	61	- .68	- .61	.71	.77
43 Best Trend Name	80	78	29	28	57	50	74	68	28	24	59	53	- .32	.00	.73	.62
44 Picture Arrangement	95	92	26	20	68	65	62	78	33	40	45	54	- .91	- .87	.49	.65
45 Arithmetic Problems	94	90	03	02	47	47	91	83	33	29	66	67	.13	.12	.88	.90
46 Identical Pictures	98	100	01	01	36	59	100	136	-01	09	60	94	.40	- .19	.75	2.68
47 Picture Group Name Selection	91	95	22	26	74	74	72	78	05	28	57	56	-1.20	-1.04	.69	.68
48 Number Classification	91	90	44	47	78	77	111	114	46	36	87	88	- .87	- .87	1.80	1.81
49 Word Exclusion	87	89	21	22	59	59	73	73	12	08	57	54	- .37	- .42	.70	.65
50 Number Relations	75	72	37	28	60	53	88	78	48	41	68	61	- .39	- .15	.93	.76
51 Word Linkage	73	76	35	35	58	61	77	74	29	38	60	59	- .49	- .46	.76	.72
52 Figure Classification	93	89	04	47	75	73	102	84	-48	37	70	62	-1.01	- .79	.96	.80
53 Class Name Selection	94	91	34	33	81	81	103	96	-08	01	70	76	-1.21	-1.13	.99	1.18
54 Necessary Arithmetic Operations	87	91	49	49	73	64	90	83	56	58	80	75	- .81	- .45	1.36	1.13
55 Verbal Analogies III	72	73	08	08	27	28	62	64	-03	-20	39	43	1.19	1.34	.43	.48
56 Remembering Classes: Members III	86	83	35	41	69	74	82	79	27	16	57	59	- .72	- .99	.70	.74

[a] Decimals have been omitted.

[b] The number of choices for each item in a test can be found in the Appendix.

items of a test is given. For some of the tests, the highest value of X_{50} and β for an item in the test could not be computed; this is the case when the biserial correlation is 1.00 or greater. When the biserial correlation is near zero, which is the case for the lowest value for a few of the tests, the X_{50} becomes very large and meaningless. As with Table 4, the tests are listed in the order of administration. An alphabetical listing of the tests with their numbers on this table is given in Appendix A. The number of choices for each item of each test can be found in Appendix B.

It should be emphasized here that these item statistics are not really appropriate for the two speeded tests, Perceptual Speed (25) and Identical Pictures (46). The low biserial correlations were for items that appeared early in the test which were answered incorrectly by only one or two subjects.

As was evident from the means of the total scores, and as can be seen from the two difficulty indices given for the items (proportion correct and X_{50}), the items generally were at the same difficulty level for boys as for girls. The obtained difficulty indices and mean scores indicate that these tests are appropriately difficult for these subjects with the exception of Verbal Analogies III which seems to be too difficult. In general, many of the tests tend to be fairly easy for fifth grade subjects.

The proportion correct indicates how many subjects responded correctly to an item. The item difficulty index X_{50} gives, in standard deviation units, the criterion score at which a subject would have a 50-50 chance of getting the item correct. For example, an X_{50} value of 1.20 for an item means that subjects with a criterion score 1.20 standard deviation units *above* the mean have a 50% chance of answering that item correctly. Subjects with a criterion score higher than this have a greater chance of answering that item correctly and subjects with a criterion score lower than this, a lesser chance. Similarly, an X_{50} value of -1.20 means that subjects with a criterion score 1.20 standard deviation *below* the mean have a 50% chance of getting that item correct; for a higher score the chance would be greater and for a lower score the chance would be less.

The two item discrimination indices, biserial correlation and β, are closely related since β is computed as a function of the biserial correlation (Baker, 1969). They are not linearly related, however. From .00 to about .30 (absolute) they are very nearly the same; beyond this, β begins to increase quite rapidly in magnitude. It may be pointed out that β is always equal to or greater (absolute) than the biserial correlation. As a general rule, .30 is often used as a lower cutting point for a desirable biserial correlation or β. For a test composed of highly homogeneous items, which is desirable for these cognitive abilities tests, the higher the discrimination indices the better.

As can be seen from Table 5, the mid-values for the biserial correlations and βs are generally quite high. The highest biserial correlation is very good for most of the tests. Some of the lowest biserial correlations are lower than is desirable—less than .30. The highest and the lowest value of the biserial correlation may not be for the same item for both boys and girls. A summary of all of the poor items, with a poor item defined as one having a β less than .30 for the correct choice, reveals that for all of the tests except the two speeded ones there is a total of 22 poor items for boys only, 36 for girls only, and 17 that were poor for both boys and girls. Thus, of the total of 1294 items, 39 of them functioned poorly for the boys and 53 of them functioned poorly for the girls. It may be pointed out here that some of these were items included in published tests.

There does not seem to be a consistent pattern in the magnitude of the mid-values of the βs for the boys as compared with the girls. For some of the tests, the βs are higher for the boys and for some of them they are higher for the girls. For the tryouts of the constructed tests, data for both boys and girls were analyzed together. If the data for boys and girls were pooled and item analyzed, the β values would probably increase for most of the items. This is also indicated by the number of items that were defined as poor ones for one of the samples and not for the other.

Almost all of the items have desirable biserial correlations and βs; this is true for approximately 95% of the items for the girls and approximately 97% of the items for the boys.

GPO 826-278-3

IV
Summary and Conclusions

The primary objective of the project entitled "A Structure of Concept Attainment Abilities" is to formulate one or more models or structures of concept attainment abilities, and to assess their consistency with actual data. One of the major steps for attaining this primary objective was taken to be the identification of reference tests for cognitive abilities. This paper describes the test development efforts for constructing 35 of the tests included in a battery of 56 tests of possible cognitive abilities. This paper also presents the item and total score statistics obtained for the 56 tests in the total battery.

Fifty-six tests were suggested by an analysis of three fairly well-known systems for defining general cognitive abilities and by a fourth schema which was suggested by this analysis. These 56 tests were administered during summer, 1970, to 172 boys and 210 girls who had just completed the fifth grade. The data were item analyzed separately for boys and for girls, using the GITAP program (Baker, 1969).

The means, standard deviations, and Hoyt reliability estimates obtained for each of the tests are presented and discussed. A summary of four different item indices—proportion correct, item-criterion biserial correlation, X_{50}, and β—obtained for the correct choice for each item of each test is presented and discussed.

Conclusions

The major conclusions drawn are:

1. With the exception of the Verbal Analogies III test, the reliability estimates obtained for these selected tests of possible cognitive abilities are sufficiently high to warrant study of the dimensionality of these tests.
2. The means of the tests and the difficulty item indices obtained indicate that these tests are of appropriate difficulty levels for these subjects except for the Verbal Analogies III test which is too difficult.
3. Almost all of the items have desirable levels of discrimination indices.

References

Baker, F. B. Fortap: a fortran test analysis package. Department of Educational Psychology, The University of Wisconsin, 1969.

Gouber, E. C. *Analogy questions*. New York: Arco Publishing Co., Inc., 1967.

Guilford, J. P. *The nature of human intelligence*. New York: McGraw-Hill, 1967.

Guttman, L. Integration of test design and analysis. In *Proceedings of the 1969 Invitational Conference on Testing Problems*. Princeton, N.J.: Educational Testing Service, 1970.

Harris, M. L. and Harris, C. W. *Newly constructed reference tests for cognitive abilities*. Wisconsin Research and Development Center for Cognitive Learning, Working Paper No. 61, in press (a).

Harris, M. L. and Harris, C. W. *Three systems of classifying cognitive abilities as bases for reference tests*. Wisconsin Research and Development Center for Cognitive Learning, Theoretical Paper No. 33, in press (b).

Thurstone, L. L. *Primary mental abilities*. Psychometric Monograph No. 1. Chicago: The University of Chicago Press, 1938.

Thurstone, L. L., and Thurstone, T. G. *Factorial studies of intelligence*. Psychometric Monograph No. 2. Chicago: The University of Chicago Press, 1941.

Appendix A
Alphabetical Listing of Tests

Arithmetic Problems (45)
Best Trend Name (43)
Best Word Class (21)
Card Rotations (18)
Circle Reasoning (10)
Class Name Selection (53)
Concealed Words (24)
Disemvowelled Words (8)
Figure Analogies (33)
Figure Classification (52)
Figure Exclusion (11)
Figure Matrix (36)
Gestalt Completion (17)
Identical Pictures (46)
Letter Classification (27)
Letter Grouping (9)
Letter Triangle (26)
Necessary Arithmetic Operations (54)
Number Class Extension (5)
Number Classification (48)
Number Exclusion (38)
Number Relations (50)
Number Series (3)
Omelet (22)
Paragraph Comprehension (14)
Perceptual Speed (25)
Picture Arrangement (44)
Picture Class Memory (28)
Picture Classification (13)
Picture Exclusion (31)
Picture Group Name Selection (47)
Picture Group Naming (23)
Picture Meaning (1)
Puzzles (29)
Remembering Classes: Members (4)
Remembering Classes: Members II (7)
Remembering Classes: Members III (56)
Remembering Classes: Names (15)
Remote Class Completion (37)
Same-Opposite (35)
Scrambled Sentences (34)
Seeing Trends (12)
Sensitivity to Order (32)
Sentence Order (39)
Spatial Relations (19)
Spelling (30)
Verbal Analogies (42)
Verbal Analogies III (55)
Verbal Classification (2)
Verbal Exclusion (20)
Vocabulary (40)
Word Exclusion (49)
Word Group Naming (16)
Word Groups (6)
Word Linkage (51)
Word Relations (41)

Apendix B
Source of Test and Number of Choices for Each Item

Test		Source	Number of Choices
1	Picture Meaning	Primary Mental Abilities Tests (1962)	4
2	Verbal Classification	Constructed[a]	3
3	Number Series	Constructed	5
4	Remembering Classes: Members	Constructed	2
5	Number Class Extension	Constructed	3
6	Word Groups	Constructed	3
7	Remembering Classes: Members II	Constructed	2
8	Disemvowelled Words	Constructed	5
9	Letter Grouping	Constructed	4
10	Circle Reasoning	Constructed	7
11	Figure Exclusion	Primary Mental Abilities Tests (1962)	4
12	Seeing Trends	Constructed	3
13	Picture Classification	Constructed	3
14	Paragraph Comprehension	Iowa Tests of Basic Skills (1964)	4
15	Remembering Classes: Names	Constructed	2
16	Word Group Naming	Constructed	Free response[b]
17	Gestalt Completion	Adapted from Gestalt Completion Test—C-1 (ETS Kit, 1962)	Free response[b]
18	Card Rotations	Card Rotations Test—S-1, Part II (ETS Kit, 1962)	2
19	Spatial Relations	Primary Mental Abilities Tests (1962)	4
20	Verbal Exclusion	Constructed	4
21	Best Word Class	Constructed	4
22	Omelet	Constructed	Free response[b]
23	Picture Group Naming	Constructed	Free response[b]
24	Concealed Words	Adapted from Concealed Words Test—Cs-2 (ETS Kit, 1962)	Free response[b]
25	Perceptual Speed	Primary Mental Abilities Tests (1962)	6
26	Letter Triangle	Constructed	3
27	Letter Classification	Constructed	3
28	Picture Class Memory	Constructed	2
29	Puzzles	Selected items from Test of Logical Ability (Hill, 1960)	2
30	Spelling	Iowa Tests of Basic Skills (1964)	5
31	Picture Exclusion	Constructed	4
32	Sensitivity to Order	Constructed	5
33	Figure Analogies	Lorge-Thorndike Intelligence Tests (1964)	5
34	Scrambled Sentences	Constructed	2

Appendix B (Continued)

Test	Source	Number of Choices
35 Same-Opposite	Constructed	2
36 Figure Matrix	Sheridan Psychological Services, Inc. (1969)	5
37 Remote Class Completion	Adapted from WADDLE Test (Warren & Davis, 1970)	Free response[b]
38 Number Exclusion	Constructed	4
39 Sentence Order	Constructed	3
40 Vocabulary	Iowa Tests of Basic Skills (1964)	4
41 Word Relations	Constructed	5
42 Verbal Analogies	Items adapted from *Analogy Questions* (Gouber, 1967)	4
43 Best Trend Name	Sheridan Psychological Services, Inc. (1969)	3
44 Picture Arrangement	Dorothy C. Adkins' adaptation of the comic strip "Louie."	Free response[b]
45 Arithmetic Problems	Adapted by and obtained from J. P. Guilford	Free response[b]
46 Identical Pictures	Identical Pictures Test—P-3, Part II (ETS Kit, 1962)	5
47 Picture Group Name Selection	Constructed	3
48 Number Classification	Constructed	5
49 Word Exclusion	Constructed	4
50 Number Relations	Constructed	4
51 Word Linkage	Constructed	3
52 Figure Classification	Lorge-Thorndike Intelligence Tests (1964)	5
53 Class Name Selection	Constructed	3
54 Necessary Arithmetic Operations	NLSMA Reports (1968)	4
55 Verbal Analogies III	Sheridan Psychological Services, Inc. (1969)	4
56 Remembering Classes: Members III	Constructed	2

[a]The tests for which the source is "Constructed" can be found in "Newly Constructed Reference Tests for Cognitive Abilities (Harris & Harris, in press). The adapted Concealed Words and Gestalt Completion tests can also be found there as can the Verbal Analogies test.

[b]Each item of this test was scored right or wrong.

GPO 826-276-2

National Evaluation Committee

Helen Bain
Immediate Past President
National Education Association

Lyle E. Bourne, Jr.
Institute for the Study of Intellectual Behavior
Universiy of Colorado

Jeanne S. Chall
Graduate School of Education
Harvard University

Francis S. Chase
Department of Education
University of Chicago

George E. Dickson
College of Education
University of Toledo

Hugh J. Scott
Superintendent of Public Schools
District of Columbia

H. Craig Sipe
Department of Instruction
State University of New York

G. Wesley Sowards
Dean of Education
Florida International University

Benton J. Underwood
Department of Psychology
Northwestern University

Robert J. Wisner
Mathematics Department
New Mexico State University

Executive Committee

William R. Bush
Director of Program Planning and Management
and Deputy Director, R & D Center

Herbert J. Klausmeier, Committee Chairman
Director, R & D Center

Wayne Otto
Principal Investigator
R & D Center

Robert G. Petzold
Professor of Music
University of Wisconsin

Richard A. Rossmiller
Professor of Educational Administration
University of Wisconsin

James E. Walter
Coordinator of Program Planning
R & D Center

Russell S. Way, ex officio
Program Administrator, Title III ESEA
Wisconsin Department of Public Instruction

Faculty of Principal Investigators

Vernon L. Allen
Professor of Psychology

Frank H. Farley
Associate Professor
Educational Psychology

Marvin J. Fruth
Associate Professor
Educational Administration

John G. Harvey
Associate Professor
Mathematics

Frank H. Hooper
Associate Professor
Child Development

Herbert J. Klausmeier
Center Director
V. A. C. Henmon Professor
Educational Psychology

Stephen J. Knezevich
Professor
Educational Administration

Joel R. Levin
Associate Professor
Educational Psychology

L. Joseph Lins
Professor
Institutional Studies

Wayne Otto
Professor
Curriculum and Instruction

Thomas A. Romberg
Associate Professor
Curriculum and Instruction

Peter A. Schreiber
Assistant Professor
English

Richard L. Venezky
Associate Professor
Computer Science

Alan M. Voelker
Assistant Professor
Curriculum and Instruction

Larry M. Wilder
Assistant Professor
Communication Arts